A Communication Device for Patients of Total and Classic Form Locked-In Syndrome

1 PROBLEM CONTEXT

Communication, in one form or another, is at the core of everything that we do as humans, with spoken language and writing being by far the most efficient and widely used media. Therefore, losing the ability to speak or write presents a massive change in the ability that a person has to communicate and forms a degradation in quality of life for that person. Hence for a long time developing novel ways to improve affected patients' abilities to communicate has been a priority in medicine and medical technology research, with technological advancements in the late 20th century and beyond driving the field forwards.

As an example, there are many neurological and spinal chord disorders that inhibit a patient's ability to talk but leave them able to fully control their eyes. For these cases methods such as eye-tracking have been developed to great effect over the previous decades, and allow patients to use gaze to locate and select characters from a computer screen, which are then recorded as text and can be passed to an artificial speech generator. Whilst incredibly successful for restoring communication abilities for many users, eye tracking does have its drawbacks. For example, latency has consistently been a major issue with these types of design: to date, most, if not all, of eye tracking approaches take a 'point-and-click' approach, where the user moves their eyes to a region representing a character or word, and fixes their gaze there for a set amount of time to select it. The first issue with this approach is that the eye tracking itself is noisy, and the noise means that there is a risk that eyes are seen to move to adjacent targets rather than the target of choice. A historical solution has been to use a dwell-time, which is a period of time that the user has to fix their gaze over a target before it is selected. Whilst this reduces the false classification frequency, it increases the latency, as now time is taken in not only moving to a point, but selecting it also. The second problem concerns the time taken to move a cursor to the target of choice. Previous research [19] found the average straight-line movement time to be around 1.4 seconds per character in their investigations, but the time can be more generally described by Fitt's Law [13] (which provides a linear relationship between target width, distance to target, and time of movement). As explained by Fitt's law, this time may be reduced by increasing the size of each target to allow the user to move more confidently and quickly towards it (as is done in [17]). However, increasing target size means that fewer options can be displayed at once, and leads to a need for a hierarchical selection process (where choosing one option leads to other options being given to the user) - thus again introducing latency issues.

In addition to the latency issues associated with eye-tracking, which are constantly being improved upon as research develops, another issue is what to do when the underlying medical condition also prevents full eye movement. These conditions are the most debilitating of all, and often go hand-in-hand with complete paralysis. The most prevalent cause of such paralysis is Locked-In Syndrome (LIS), which has a prevalence expected to be about 1 in 10,000 [8], and is brought on by a wide variety of causes, most often relating to traumatic injury or strokes. There are three categories of severity, *Incomplete, Classic,* and *Total* LIS, with incomplete LIS leaving partial muscle control (often including vertical eye movements), and classic and total LIS typically being more acute and losing even the control of eye muscles [15]. Therefore, the eye-tracking methods as described above become unusable for patients of classic and total LIS. What is more, the above methods rely on greater variation of eye movements than only vertical, and so even for the incomplete LIS, eye-tracking alone is currently insufficient to deal with this issue[1].

There have in the past been user-specific cases where alternatives to eye-tracking have been used for text generation - a famous example is with Stephen Hawking, who used eye-tracking software for many years after suffering speech loss, but eventually had to move to cheek sensors instead when his eyelids drooped too much and prevented reading of his eye movements [9]. However, these specific cases are not scalable, as they involve a large amount of research and development for each case and therefore are not suitable for the thousands of people who suffer LIS. What is more, the extent of muscle control loss brought on by LIS means that whilst some patients might retain control of other part of their bodies for some time, it is likely that this would only be temporary, and it is not generalisable to other patients. The device designed in this paper would act to bridge this gap, and provide a robust and effective means of generating text without the need to move any muscles. Inspired by previous work involving the use of electroencephalograms (EEGs) to translate neural activity into classification choices [11, 14, 18], this device's proposed functionality would be to analyse neural signals with complex machine learning signal processing to provide the text that the patient intended to produce.

1.1 Influential Factors

In order to be classifiable by a machine learning model the neural signals recorded must have low signal to noise ratio (SNR). It must also be possible to distinguish up to 41 characters/intentions (26 ASCII letters, 10 numerical characters, an activation/termination symbol, a space, a comma, a full-stop, and an apostrophe) from the neural signals, and so they must have the flexibility to cater to this. A method that has recently been the focus of a large body of research in brain-computer interfaced (BCI) design is therefore to take recordings from the motor cortex, as it has a very large diversity of activation patterns brought on by the diversity of motor signals that human muscles produce. This means that if there are a enough sensors then there is sufficient temporospatial variability in the signal vectors for a machine learning model to learn to distinguish between 41 classes. Patients of LIS are able to generate motor signals even when their motor function is completely impaired, as the motor cortex generates the *intention* to move a muscle, which is then passed onto the lower regions of the motor system hierarchy (which are in some way damaged in LIS patients) [18]. Therefore, by imagining performing a highly dexterous task such as producing a handwritten character, the user can generate the correct corresponding motor cortex signals without being able to actively execute the motion.

In addition, the product is intended to be used in a continuous setting, where it is able to be recording neural signals and analysing them constantly even when not in active use. This way when a known activation symbol (consisting of a neural signal representing a motor intention that is not temporospatially similar to any other character) is given from the user, the system is activated and ready to generate text from the subsequent data.

[1]Previous methods for communicating with patients of incomplete LIS have involved the AEIOU board, where a patient specifies one of 5 lines using vertical eye movements, and then moves down that line to select a character. This approach is extremely slow however, and currently requires a human translator to decode the message [15].

MLMI 10, Designing Intelligent Interactive Systems,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Similarly, when the user is finished generating text, the same symbol can be repeated as a termination symbol to switch the system back into hibernation mode, where the predictions are not printed to a screen.

1.2 Solution-Neutral Problem Statement

Based on the context as described in the section above, the solution-neutral problem statement is as follows:

Design a non-invasive device that can be used to restore communication abilities to patients of total and classic locked-in syndrome by generating text at a rate comparable to average human typing speed.

2 REQUIREMENTS SPECIFICATION

2.1 Requirements Elicitation

The target audience for this project make the requirements elicitation naturally different to other products' as, for example, focus groups and surveys of severely affected LIS patients are difficult to conduct. Nonetheless we must consider what the preferences of the patients may be, and how this product can be made as user-friendly as possible. There are two approaches that can be taken to inform this; talking with close friends and family of LIS patients who are familiar with the realities of living with the condition, and consulting medical professionals with specialisms in neural injuries who can advise on the practicalities of the designs such as optimal sensor placements. To this effect, the requirements gathering stage of the design process will involve interviewing a small number of people (as using large numbers would require interviewing people without any exposure to LIS) with relevant medical expertise or experience caring for a LIS patient. In a small number of cases it may even be possible to interview LIS patients themselves, using a communication medium such as one of the current methods discussed in the problem context, and simple questions that do not require involved answers in order to cater for the difficulty of communicating in this way.

2.2 Function Analysis

The function analysis for the neural text generator can be seen as a Function Analysis and System Technique (FAST) diagram in Figure 2. From left to right it builds up from the most basic functionality of the device to greater detail into how functions can be executed (i.e. moving left to right in the tree explains how the root functions are executed, and moving right to left answers the question of why the leaf functions are necessary). The prediction function in the execution stage of the FAST diagram involves a pre-trained machine learning model, and the language model acts as an additional probabilistic measure to help reduce errors from misclassifications in the character classification from the encoded neural signal.

		Process	Performance	Safety	Cost	Documentation
Product in Use	Operation	Operation Process	Operation Performance	Operation Safety	Operation Cost	Operation Documentation
	Maintenance	Maintenance Process	Maintenance Performance	Maintenance Safety	Maintenance Cost	Maintenance Documentation
	Disposal	Disposal Process	Disposal Performance	Disposal Safety	Disposal Cost	Disposal Documentation
Product Design/ Manufacture/ Supply	Design	Design Process	Design Performance	Design Safety	Design Cost	Design Documentation
	Manufacture	Manufacturing Process	Manufacturing Performance	Manufacturing Safety	Manufacturing Cost	Manufacturing Documentation
	Distribution	Distribution Process	Distribution Performance	Distribution Safety	Distribution Cost	Distribution Documentation
	Installation	Installation Process	Installation Performance	Installation Safety	Installation Cost	Installation Documentation

Figure 1: A requirements matrix used as a guide to outline the requirements that are most relevant the for the specification for this device. The boxes highlighted in green are expanded in the requirements list below.

2.3 Requirements Matrix Checklist

The matrix in Figure 1 highlights the most relevant specification areas to this device in order to guide the more detailed elicitation of requirements below.

The specifications can therefore be broken down into more detail as shown in the checklist below. The source is given for each element of the list, where BJK indicates that the requirement originates as a crucial design feature rather than a legal requirement or government regulation.

Operation Process:

- A) Be solely electrically powered (source: BJK)
- B) Be able to override on/off functions via physical switch (source: BJK)
- C) Does not output any form of media when in standby mode (source: BJK)
- D) Output is generated sequence when in active mode (source: BJK)
- E) In case of model error of not being able to access input readings, print error statement on screen (source: BJK)

Operation Performance:

- A) Non-invasive sensors (source: BJK)
- B) Sensors to be fixed within a fastenable accessory in order to remain within ±5mm of the desired location on the head (source: BJK)
- C) Able to operate without more than 30 hours of training (source: BJK)
- D) Not to have visible naked sensors (source: BJK)
- E) Whole device to weigh <3 kg (source: BJK)
- F) Any wearable part of the device to weight <0.5 kg (source: BJK)
- G) Operating temperature range to be between -40 and 85 degrees Celsius (source: operating temperature of microchips [3])
- H) Processing unit of the device to be smaller than $5cmx4cmx10cm$ (source: BJK)
- I) Less than 10% character error rate (CER) when functioning (source: BJK)
- J) Greater than an 80 characters per minute (CPM) text generation rate (comparable efficiency to eye-tracker methods, equivalent to approximately 20 words produced a minute) (source: BJK)
- K) Any graphic display to have an area greater than 3x5cm and less than 30x20cm (source: BJK)
- L) Any audio output device to have volume range 10-85dB (source: BJK)
- M) If on battery power, have battery life >4 hours (source: BJK)

Operation Safety:

- A) The device and all materials that are part of it must be approved by the Medicines and Healthcare products Regulatory Agency for sale in the UK (source: MHRA)
- B) The device and all materials that are part of it must be approved under Regulation (EU) 2017/745 on Medical Devices for sale in European Union countries (source: EUR-Lex EU Regulation Database)
- C) The device and all materials that are part of it must be approved by the Center for Devices and Radiological Health (CDRH) for sale in the United States (source: U.S. Food and Drug Administration FDA)
- D) The device must comply with electrical medical device safety standards, EN 60601, in order to be sold in the UK (source: BSI British Standards), for example the entire device must be earthed and electrically insulated
- E) User specific data collected to be stored on device and not externally accessible unless explicit permission given by the

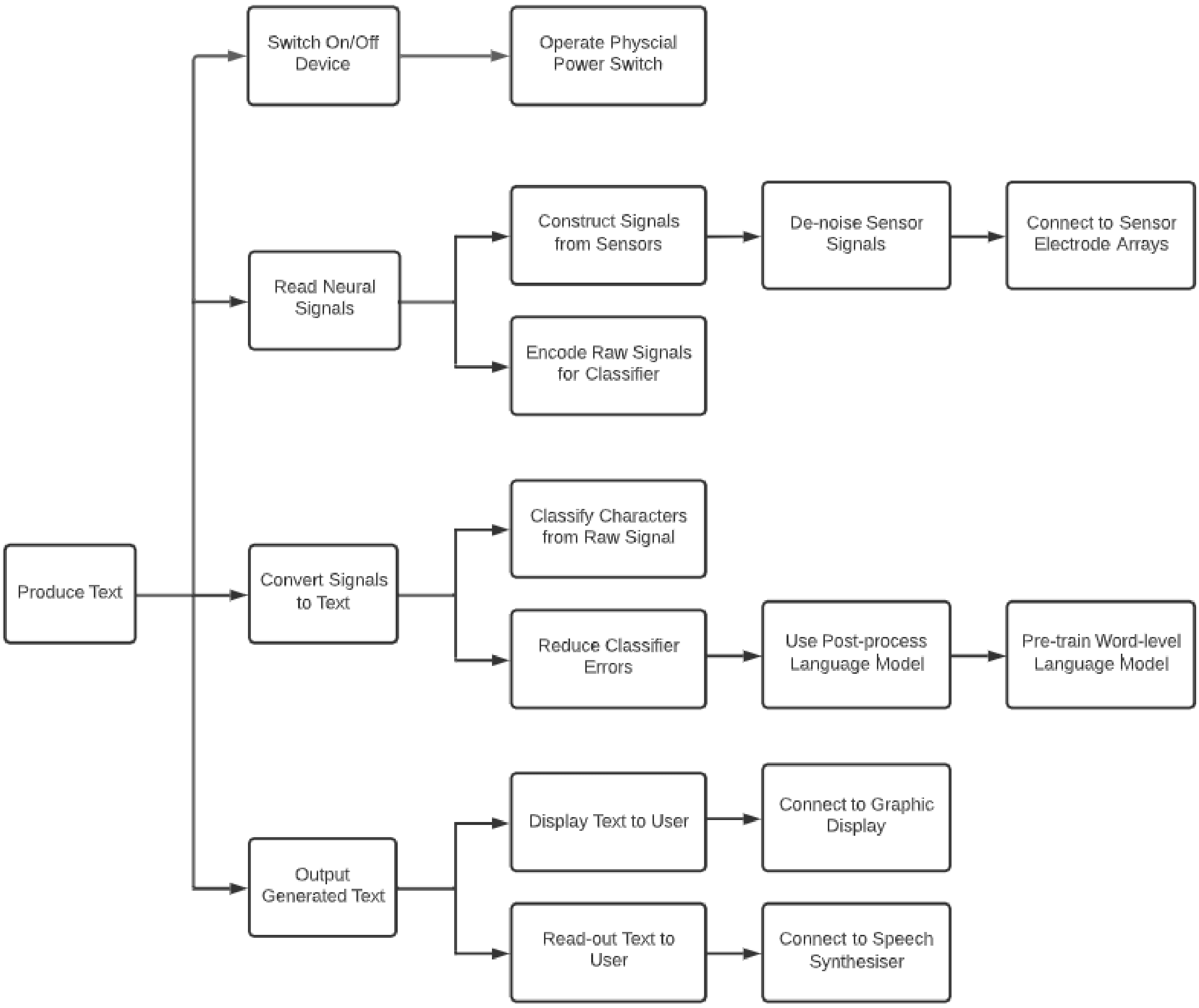

Figure 2: FAST diagram for the proposed text generation device. Moving left to right signifies extra detail into how the device carries out functions, through basic and secondary functions.

user, in order to comply with data protection (source: Data Protection Act 2018)

- F) If permission granted and user data to be exported to an external device, storage and use of this data must comply with data protection laws (source: Data Protection Act 2018)

Operation Documentation:

- A) Provide a manual of operating instructions, including a visual reference for each of the characters and their meaning (source: BJK)

Maintenance Process:

- A) Automatically re-align the users writing style with the original classifier's training dataset each day[2]

Maintenance Performance:

- A) Have a warranty and support period of 10 years (source: BJK)
- B) Provide next-day support where reasonable in case of device malfunction (source: BJK)
- C) Require minimal retraining of the classifier (required due to neural plasticity) with no more frequent than monthly re-calibration (defined specifically to be 30 days) (source: BJK)

[2]This can be done for example via unsupervised learning techniques, as described in [18], where a transformation is learned between today's predicted character and the corresponding character from the previous day. This transformation is then applied to that character for the current day to limit degradation of performance due to neural plasticity.

Manufacturing Process:

- A) Pass/Fail test every product before mass production phase to ensure the device functions as expected (source: BJK)
- B) Pass/Fail test every 1 in 10 products when in mass production in order to validate that the devices are functioning as expected (source: BJK)

Manufacturing Cost:

- A) Cost less than £5,000 per unit to manufacture before mass production (source: BJK)
- B) Cost less than £2,000 per unit to manufacture once in mass production (source: BJK)
- C) Keep device assembly costs below £10 per device once in mass production (source: BJK)
- D) Maintain a 10% held-back store of components for use in maintenance and repairs of sold devices (source: BJK)

Installation Process:

- A) Carry out basic functionality checking when fitting and calibrating devices (source: BJK)

Installation Performance:

- A) Any non-wearable section of the device (for example an output device or processing unit) to be fixable to a wheelchair without any tools or expertise needed (source: BJK)

- B) Trained individual to aid in calibrating the device to the user upon first installation (devices require calibration as different users have differing neural signals) (source: BJK)

3 CONCEPTUAL DESIGN

3.1 Function Structure

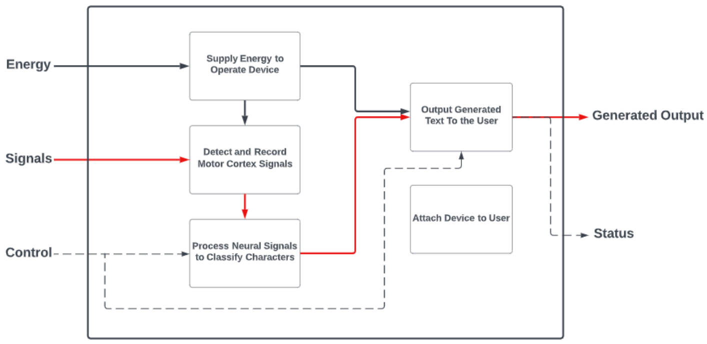

Figure 3: Diagram representing the sub-function structure of the device. These are the core functions that need to be addressed in the conceptual design.

Figure 3 shows the core functions which the device needs to perform in order to operate in the manner laid out in the FAST diagram of Figure 2. The red lines represent the core pathway from input to output of the device, meaning the neural signals through to the generated text in this case, and the dashed and black lines represent the control signals and the energy flow respectively (as discussed in Section 1, the control needs to be able to activate and terminate the process upon instruction from the user, and will thus also be represented by an invented character generated by a motor cortex signal). From here we are able to propose solutions for each sub-function in order to build up a workable design for how the device will function, and this is shown in the section below.

3.2 Morphological Chart and Generated Designs

Figure 4 shows three potential solutions for each of the sub-functions specified in the function structure diagram. In the table the purple circles and yellow triangles give two different paths through the chart and represent alternative designs for the device that perform the functions detailed in the FAST diagram in Figure 2. Note that in a few cases we are able to select multiple solutions for the same sub-function if it will improve the design - for example, one path uses both a transformer based sequence-to-sequence model for character classification followed by a pre-trained language model, which is used to assign probability scores to the classified characters to correct potential misclassifications and improve the character error rate.

Concept 1: EEG Cap Device

The first design uses EEG sensors placed on the surface of the head over the motor cortex to read action potentials of the motor system. In this way, even if motor actions are not followed through due to nervous system damage, the intention to move in that way can be recorded and used to classify the intended action. At each point in time the action potentials are then encoded and passed through a sequence-to-sequence transformer model which classifies the characters. As mentioned above, a secondary pre-trained language model can then be used to improve accuracy by scoring the classifier predictions and making corrections when improbable characters are classified, which would likely be the result of overly noisy EEG signals for that character. The device is powered both by batteries to improve portability and mains power whilst the battery is recharging, and uses a separate display and speaker to output generated text and speech from the processing unit. Finally, this concept would use a cap-like design to encase the sensors in order to keep them in place and to improve the aesthetic factor for the user.

Concept 2: sEMG Watch

The second concept instead uses surface electromyography (sEMG) to sense motor unit action potentials from nerves in the arm. The arm is chosen as the philosophy behind this device is that imagined hand-written characters still activate the motor system even after full body paralysis by LIS [18], and the motor pathway activated by hand writing mostly involves muscles in the arm. This device would use rechargeable batteries, but would not be able to recharge whilst in use due to the location of the device on the wrist. In the interests of reducing processing unit size, it would also need to use less memory than the above, and therefore a slightly less complex classification model with fewer parameters such as an RNN model has been chosen. Once again, a pre-trained language model can be combined to reduce the character error rate post-classification. The final difference is in how the output text is provided to the user. As a main reason for generating this design idea is to keep the device small and easily wearable, it makes sense to keep the outputs small and lightweight also. Therefore a small screen that is able to comfortably fit on the wrist of users is proposed, and can display the generated text. One drawback of this however is that the watch would not have room to fit in a speaker, and so there would not be built-in spoken output for this design.

3.3 Concept Evaluation

In order to quantitatively compare the two design concepts, a rough importance weighting can be assigned to a set of characteristics for each, and the total weighted value of each design generated. This is shown in Figure 5, and shows that the EEG cap device described in concept 1 above has higher value than the sEMG watch device described in concept 2. Whilst these values are rough estimates, they are still able to highlight the relative advantages and disadvantages of each design and assist in the choice of design going forwards.

3.4 Chosen Design and Function Model

Based upon the evaluation table in Figure 5, the chosen design is therefore the EEG-based device. This is mainly driven by three factors; firstly, the larger output screen makes the generated text more easily read, secondly, it has speech output which is useful when the user wishes to communicate via speech, and lastly, EEG signals have the advantage of higher SNR and temporospatial range than sEMG signals from arm nerves, which make classification easier and likely therefore more accurate. The disadvantage of being larger and less portable is outweighed by these advantages once we take into consideration that the users already have to use wheelchairs due to their condition, and therefore there is always a frame on which to attach the device.

Figure 6 therefore shows the function model for the chosen device, including the inputs, details of the connectivity between the processing stages, and the output media. At a fundamental level, the inputs are the raw EEG signals which are processed within the device, and the output is the generated text on a display and the optional sound output to read aloud the text on the screen. The user is able to choose whether they wish to power the device using mains or battery power, for example depending on whether they are away from a power source temporarily.

4 EMBODIMENT DESIGN

4.1 Design Inspiration:

As a first step in the embodiment design it is useful to take inspiration from previous design suggestions. Figure 7 shows the electrode array set-up from [18], which discusses a method for using intracortical arrays to classify text from motor cortex

Solution: / ***Sub-function:***	**1**	**2**	**3**
Supply energy to operate device	Mains power ●	Battery Power (rechargeable) ● ▲	Battery Power (non-rechargeable/replaceable)
Detect and record motor cortex neural signals	Intra-cortical (invasive) EEG sensor array to record action potentials at precise points in the motor cortex	Surface EEG (non-invasive) sensor array to record action potentials over the near regions of the motor cortex ●	Surface Electromyography (sEMG) sensors to detect low SNR motor unit action potentials in the nervous system ▲
Process neural signals to classify characters	Use of a RNN model for sequence-to-sequence text generation ▲	Use of a transformer based model for sequence-to-sequence text generation ●	Use of language model post prediction to reduce misclassification error ● ▲
Output generated text to the user	Output generated sequence of characters as text on a computer display that is connected but separate to the device ●	Output generated sequence of characters as speech and communicate via a connected speaker ●	Output generated sequence of characters as text on a small display mounted on the surface of the device ▲
Attach to user's body and contain both sensors and processing unit	Keep head sensors exposed (not in a cap) and connect them to a separate processing unit via wires	Use baseball cap design with fasteners on the side to keep in place on the user's head ●	Use a watch-like wrist-based device with an in-built sEMG sensor array ▲

Figure 4: Morphological chart showing the break-down of the sub functions of the device and possible three possible solutions for each.

Criteria:	Weighting:	Concept 1: EEG Cap Device		Concept 2: sEMG Cap Device	
		Value:	Weighted Value:	Value:	Weighted Value:
Classification Accuracy	10	8	80	5	50
Battery Life	7	9	63	6	42
Portability	4	6	24	9	36
Ease of Use (i.e. reading output)	10	9	90	5	50
Appearance/Discreetness	5	7	35	9	45
Comfort	10	8	80	7	70
Weight (wearable part)	8	8	64	9	72
Weight (non-wearable part)	2	5	10	10	20
		Total:	446	Total:	385

Figure 5: Evaluation of the two concepts using rough estimates of importance weighting for the different features of the devices.

signals of users with no other means of communication. This design is very relevant in that it focuses on the motor cortex and demonstrates in the paper the proof of concept that motor cortex signals can be used to classify text with low error rate and high characters per minute.

However, it is a strong requirement for this device that the sensor arrays be extracortical as to allow the user to remove the device easily and not require a surgical procedure. Therefore, Figure 8 gives an idea of how extracortical sensors can be used to record neural signals, with the right of the figure showing example EEG recordings for the sensor placements on the left. It is important to note that action potentials of the surface electrode recordings are approximately an order of magnitude smaller than those of the intracortical sensors due to volume dispersion of the potentials within the skull, causing a lower SNR. Therefore the additional step of denoising the signal prior to encoding needs to be added when using extracortical sensors. Finally, this plot shows the EEG sensors to be spread almost uniformly over the surface of the top of the head, but as discussed above, our design focuses on the motor cortex, and hence there is a concentration of these sensors in the anterior-central region of the brain. Figure 8 shows this region to have the highest recorded action potentials too[3], so this is expected to improve the quality of the signals.

4.2 Physical Design:

The central part of the device design is how to encase the surface EEG sensors in such a way that they will not move around on the head more than $\pm 5mm$ (in accordance with the requirements specification) and will keep the SNR as high as possible. We also wish to make the device discrete and comfortable to wear for the user. Therefore the decided design from the morphological chart resembles a baseball cap, where the EEG sensor arrays are embedded in the inside of the hat. Either side of the cap will also include methods of fastening to the user's ear in order to make sure the sensors are correctly aligned in use and reduce movement while wearing. The inspiration for these fasteners comes from earphones that are designed to stay in-place within the ear during intense exercise whilst retaining comfort. For example, a commonly used design is as in Figure 9 (Left).

For the same sensor alignment reasons, it is also important that the cap and its materials are rigid, so that sensor positions do not change when small forces are exerted on it. As discussed in the above section, the electrode arrays should be concentrated over the motor cortex in a way similar to Figure 7, whilst keeping them on the surface of the head. There are thus two electrode arrays, each $5cm$ by $5cm$ (they have to be bigger than the $4mm$ size of the intracortical array in 7 to account for signal dispersion), and containing in total 200 electrodes.

In order to provide comfort, these arrays are to be embedded in a wool material cap, but with a thin frame that the sensor arrays attach to that is made from a non-brittle but rigid material such as titanium to remove elasticity from the wool. Titanium is an expensive material, but one that is known for its toughness and durability, without being brittle. It is also commonly used in

[3] As expected, as one of the main reasons for choosing to focus on the motor cortex is because of its size and resultant magnitude and temporospatial range of action potentials. [12].

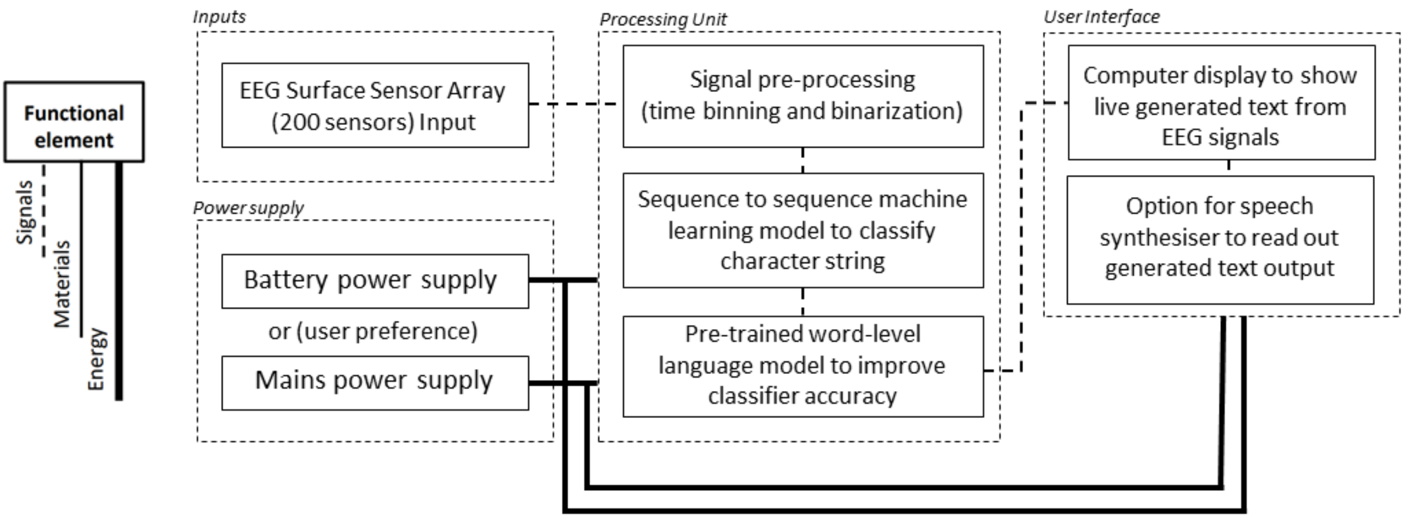

Figure 6: A diagramatic layout of how the device will be connected and perform.

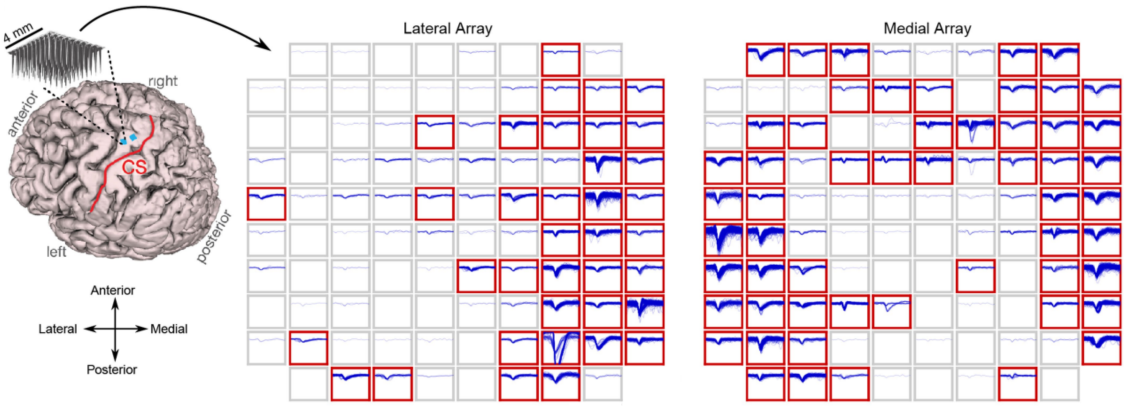

Figure 7: A previously used intracortical EEG design with 192 electrodes in two small electrode arrays (medial and lateral) which have been surgically inserted into a user's motor cortex. The red highlighted cells indicate that the sensor in that location has detected an action potential with firing frequency above the threshold (set at 2Hz for this figure), and thus the binary value for that cell is 1. This figure is adapted from [18].

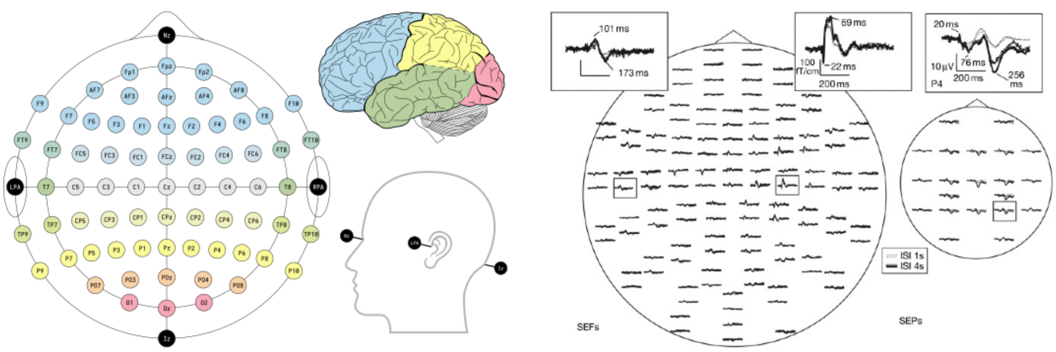

Figure 8: (Left) An alternative to intracortical electrode arrays, using instead sensors on the surface of the head. This example has sensors spread across the whole top of the head, whereas the device in this proposal would have sensors concentrated over the motor cortex as in Figure 7 (middle centre of the blue region in this plot). (Right) Example action potentials that are detectable by the surface sensors. The amplitudes are an order of magnitude less than intracortical sensor values due to dispersion of the potentials through the skull. Therefore sensor signal denoising must be performed prior to the binary encoding process for the classifier. Figures adapted from [7] and [10].

Figure 9: (Left) Example design for in-ear fasteners to keep the cap containing the EEG sensors in place. Figure taken from an earphone manufacturer on Amazon [1]. (Right) Elastic silicon straps for securing the processor case to a wheelchair. Figure taken from a manufacturer on Amazon [5].

medical contexts due to its low health risk. This can be seen in the rough (not-to-scale) sketch of Figure 10, alongside details of the sensor placements and processing unit attachment (see below for discussion).

For comfort and discreetness, the processing unit and output devices will not be inside the cap but instead connecting to it via wiring. The processing device will also have the capability to be plugged in to mains electricity, but will include an additional rechargeable battery power source to allow the device to be used when away from mains power. With respect to the design of the processing unit, it needs to be small enough to not be cumbersome, and also made from a durable material to ensure long product life-span even with constant use. However, the choice of material is not as critical for the processor case as it is for the hat, as the processor is not worn directly on the body or subject to the same amounts of force. Therefore, to keep costs lower, a cheaper material such as reinforced plastic may be used. The case will also have a means of discreetly attaching to a wheel-chair so that it does not need carrying. For this purpose, silicon elastic straps that are attached to the processor case can be used, as shown in Figure 9 (Right).

Another design consideration is for the output devices. There are two output devices, a screen for displaying the generated text, and a speaker to connect to the speech generation system. For the screen, it needs only to display sequences of characters and must be light and small enough to make it easy to attach to a wheelchair frame without modifications to the wheelchair. A standard 14-inch screen as used in laptops is therefore suitable for this purpose, being light and compact whilst large enough to make reading text from it easy. For the speaker, a low wattage single-driver speaker is sufficient as there is not the requirement to go to high volumes or to have a very long frequency range. The benefits of using a small low-energy speaker are that they are cheaper, keeping costs lower, and the low wattage means that smaller capacity batteries can be used to power the device.

The final considerations are for the battery and wiring. The battery is intended to be used when there is no availability of mains power, and thus will have to have enough capacity to power the device for at least 4 hours when fully charged. The fact that they must be rechargeable means that lithium ion batteries are a sensible choice, owing to their robustness to repeated re-charging and power efficiency [6]. A rechargeable lithium ion 5.2Ah and 10V battery as used to power laptop devices is therefore suitable for this device, and will be enclosed within the case for the processing unit. When the device is under mains power, the rechargeable battery is recharged simultaneously. The wiring must all be well insulated, in particular the wiring between the device and mains power. All of the wires (between the processing unit and headpiece, and between the processing unit and outputs) are to be given strong rubber casing also, in order to improve device robustness and longevity.

4.3 Cost of Materials and Manufacture:

An estimated break-down of material and manufacture costs can be summarised as follows for each device [4]:

(1) Wool baseball cap with embedded thin titanium frame and in-ear fastenings: £30
(2) Two EEG Sensor arrays[5]: £2, 000
(3) 14 inch monitor screen: £100
(4) Single-driver 10W speaker: £20
(5) 5.2Ah Rechargeable Lithium-ion battery: £50
(6) Arduino processing unit, such as the Portenta H7 Module with Python compatibility [3]: £80
(7) Connecting cables and charging cable: £10

Therefore, the overall material cost is expected to be around £2,290 per device before mass production. This does not include manufacturing costs as these are difficult to predict accurately, but expected to be negligible in comparison to material cost once mass production is used.

4.4 Details on Computation

As discussed above, the processing unit consists of 4 sub-processes: de-noising the EEG sensor inputs; binary encoding the de-noised inputs based upon whether the signal within each 5*ms* time-bin passes a threshold frequency; passing the encoded vectors into a sequence-to-sequence model for classification of characters; and using a pre-trained language model to correct the produced character sequences. This section looks in more detail at the sequence-to-sequence classifier.

A transformer is proposed as a means of classifying the text string due its flexibility of the inputs. Specifically, there is a lot of variation in the length of time it takes to write different characters, and therefore a variable number of historical encoded input vectors are relevant for each character. The way in which transformers can pay attention to inputs a lot further in the past than simpler models without suffering catastrophic forgetting makes them well adapted for this task. It is also expected that there will be a lot of sparsity in the inputs due to the sensors being spread over quite a wide area (neural signals in the brain tend to fire locally due to the architecture of neurons, and thus characters will likely activate clusters of sensors at a time), and therefore the ability of a transformer to assign attention intelligently to the inputs makes it suitable for this purpose. Finally, previously predicted characters contain information on what future characters are likely to be, and so transformers' ability to look back (but not forwards) in the outputs to gain extra contextual information is also useful. The proposed model is therefore similar in structure to the transformer paper in [16], but with a single multi-head attention layer in both the encoder and decoder in order to keep lower the number of parameters.

5 RISK ASSESSMENT

In order to gauge risks in the the users' operation of the device, a Structured What-If Technique (SWIFT) table is produced. Each identified risk is associated to a region within the risk matrix shown in Figure 11 which outlines the severity of that risk, with the predicted risk without recommendations shown in brackets. Risks with low severity do not risk harm to any individual and are easily fixed. In contrast, higher severity risks require safeguards to mitigate the likelihood of them happening.

A simple and brief risk assessment framework is used as this device is not a high risk device (i.e. it is not used in high risk situations, does not carry out any functions that are likely to cause harm, and whilst being a medical device, has no parts that are inserted into the body or deliver medication). The risk assessment is also brief because many extremely improbable risks that are common to most hardware devices are not included.

6 VERIFICATION AND VALIDATION

6.1 Verification Cross-Reference Matrix

Table 2 outlines how to verify each of the requirements laid out in Section 2.2. *Inspection* is the examination of the device without

[4]Note that this is not considering the effects of mass production in reducing costs, and therefore this estimate is for pre-mass production.

[5]This is the hardest part to price, as they would need to be custom made and prices vary a lot based on design. However, this estimate is given based on costs of similar, designs such as from OpenBCI [2].

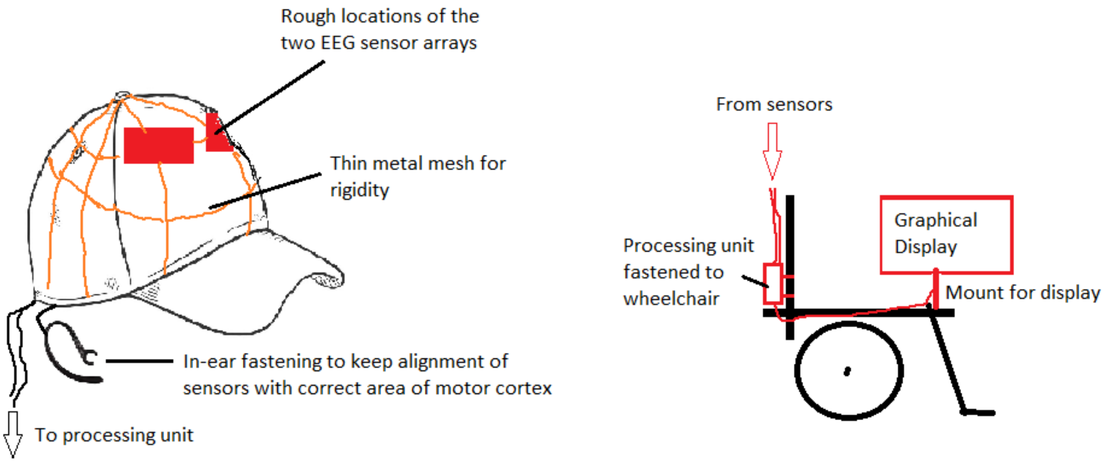

Figure 10: A rough, not-to-scale sketch giving an idea of the what the design would look like and how it would attach to a wheelchair.

Category	Risk Lvl	What-If	Consequence	Recommendation
Electrical Fault	4D (5D)	Cable breaking	Device malfunction and risk of minor electric shock	Ensure cables are reinforced with an insulating material and well earthed.
Electrical Fault	1C (2B)	Severe battery over-heating or explosion	Device malfunction and risk of minor burns	Ensure batteries are not damaged at device assembly stage and make sure battery casing is robust to withstand battery explosion.
Software Error	5E (5E)	Classification model error	Output text contains incorrect information	Pre-trained language model used to reduce the frequency of falsely classified characters. There will always be frequent character errors however due to the nature of the task, but the language model can ensure that these errors are very unlikely to be severe, such as incorrectly generating offensive language.
Physical Damage	2D (3D)	Processing unit or output devices' casing breaks	Device at risk of malfunctioning and possible sharp edges can harm user	Ensure fasteners are well connected to both device and the wheelchair upon installation so that the processing unit and outputs do not fall off. Ensure that they are made from robust materials and check for damage pre-installation.
Physical Damage	3E (3E)	Tear in cap fabric and sensor exposure	Risk of sensor damage and device malfunction	Ensure that it is functioning correctly at installation. The warranty scheme then allows for the repair or replacement of parts if they are damaged by no fault of the user within 10 years of purchase.
Electrical Fault	2E (2E)	Fault in graphic display	Generated output text not correctly displayed	Ensure that it is functioning correctly at installation. The warranty scheme then allows for the repair or replacement of parts if they are damaged by no fault of the user within 10 years of purchase.
Electrical Fault	2E (2E)	Fault in speech generator	Generated output not correctly spoken	Ensure that it is functioning correctly at installation. The warranty scheme then allows for the repair or replacement of parts if they are damaged by no fault of the user within 10 years of purchase.
Electrical Fault	2E (3E)	Device heats up over normal operating ranges	Risk of device malfunction, extremely improbable risk of burning	Carry out quality checks of circuit boards and components and ensure device passes electrical device medical safety standards EN 60601 prior to distribution.
Market Risk	N/A	Competitor device introduced at same time as ours	Risk of low uptake of this device and/or shortened product life-cycle	Ensure that users' opinions are actively sought out so that we can be confident we are designing a desirable and relevant product for them.

Table 1: SWIFT analysis of the causes of the most likely risks in the device, and how their risks can be mitigated. The brackets in the risk level show what the risk would be if recommendations were not carried out.

seeing it in use, *analysis* is a statistical treatment involving a user group and pre-defined tests and models, *demonstration* refers to simply observing the device functioning under usual conditions, and *testing* refers to carrying out accurate tests on pre-defined tasks in order to evaluate specific functions.

6.2 Validation

Validation concerns the assurance that the device meets customer needs satisfactorily, which means we want to check the usability and practicality of the device from the users' perspective. The intended users of this device will all be patients of severe neural injury or disease that leaves them heavily paralysed, with specific focus given throughout this design to patients of total and classic form LIS, who notably do not have control over a large range of motion (often all motion) of their eyes. Therefore, to them usability means that they are able to operate the device once installed without requiring any input other than their intent to move their writing arm to write certain characters.

Risk probability	Risk severity				
	Catastrophic A	Hazardous B	Major C	Minor D	Negligible E
Frequent 5	5A	5B	5C	5D	5E
Occasional 4	4A	4B	4C	4D	4E
Remote 3	3A	3B	3C	3D	3E
Improbable 2	2A	2B	2C	2D	2E
Extremely improbable 1	1A	1B	1C	1D	1E

Figure 11: Risk matrix outlining the severity of the risks highlighted by the SWIFT table in Table 1. Figure taken from [4].

A simple and effective test to check that the device can do this is to demonstrate it working directly on these users. Of course the users will all be slightly different, for example with varying neural signals and plasticity, so performance would need to be validated on a large enough group of users (>10) to confidently validate whether the device meets its requirements.

From a user's perspective, the device must also be comfortable, visually appealing, and simple to learn to use. These factors can be validated by asking the users in the user group to comment on how they would quantitatively rate these factors out of 10, and asking for feedback on why they gave the score they did. These factors are particularly important given that they will spend so long wearing and using the device upon distribution, so success in this context would have to mean that they agree on a high rating for all three factors.

7 SUMMARY

This design proposal provides a detailed outline for developing a device that translates an LIS patient's handwriting motor intentions into a string of characters, to produce text in a fast and accurate manner. The device will be portable and simple to use, with both mains charging and battery functionality in order to be flexible to the requirements of the user's everyday life. The intention is that this device will allow the users more communication ability than has been previously possible, so that they can live a higher quality of life and are better able to leave the confines of the hospital as they would no longer require a trained interpreter (for example trained to use AEIOU boards) to communicate.

The settled-on design uses an EEG array of 200 electrodes contained within a wearable cap that is fixed in place to minimise electrode disalignment. New advancements in machine learning algorithm efficiency and complexity means that much more noise can be present in the signal before large decreases in classification accuracy are observed, which means that the non-invasive nature of the surface electrodes and small-scale slippage of sensors no longer present the major challenges that they have done in the past.

The machine learning framework used to build predictions is split into three parts:

(1) De-noising the raw signals from the EEG sensors.
(2) Encoding of the de-noised potentials by time-binning the signals every $5ms$ and using a binary encoding for each of the 200 sensors representing whether they are greater than a certain threshold frequency or not (e.g. $2Hz$ suggested in [18]) - this produces a 200 x 1 binary valued vector every $5ms$.
(3) Passing these vector encodings into a sequence-to-sequence classification model to translate them into strings of characters (41 characters available, see section 1.1) representing the intended text. 1 of the characters is an action command, that either activates the text output feature (if previously in hibernation) or deactivates it (if previously active).
(4) A probabilistic language model helps to reduce misclassifications of the predictive model by calculating the probability that the currently generated word represents each word in a dictionary as each character is generated, and weights character selections based upon its prediction of the most likely word being written (in the same way that predictive text on phones does).

Due to plasticity in the brain it is expected that neural signals will change for the same characters over time, meaning that performance will gradually decrease if re-alignment is not performed. This re-alignment will be carried out in an unsupervised statistical manner every 24 hours to mitigate the impacts of this. However, this is not expected to be a perfect fix, and so there will need to be routine maintenance to the system that involves complete model retraining, predicted to be 1 month (30 days). Whilst this will take time and effort for the user, it is believed that the benefits of this far easier and more flexible means of communication will outweigh any occasional annoyances and will therefore not be an issue for the user.

REFERENCES

[1] 2022. Futheda 10pcs bluetooth headset earpiece receiver clip. https://www.amazon.co.uk/Futheda-Bluetooth-Earpiece-Receiver-Replacement/
[2] 2022. OpenBCI featured products. https://shop.openbci.com/collections/frontpage/eeg
[3] 2022. Portenta H7. http://store.arduino.cc/products/portenta-h7
[4] 2022. Safety Management. https://www.icao.int/safety/SafetyManagement/Pages/default.aspx
[5] 2022. Silicon Mountain Bike Flashlight Elastic Accerssories. https://www.amazon.com/Silicon-Mountain-Flashlight-Elastic-Accessories/
[6] 2022. What are the benefits of lithium-ion batteries? https://www.tennantco.com/en_ca/blog/2019/11/lithium-ion-battery-benefits.html
[7] George H. Klem, Hans Lüders, Herbert H. Jasper, and Christian Erich Elger. 1999. The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalography and clinical neurophysiology. Supplement* 52 (1999), 3–6.
[8] Roy Kohnen, Jan Lavrijsen, J Bor, and Raymond Koopmans. 2013. The prevalence and characteristics of patients with classic locked-in syndrome in Dutch nursing homes. *Journal of neurology* 260 (01 2013), 1527–1534. https://doi.org/10.1007/s00415-012-6821-y
[9] Joao Medeiros. 2015. How intel gave Stephen Hawking a voice. https://www.wired.com/2015/01/intel-gave-stephen-hawking-voice/
[10] J.P. Mäkelä. 2014. 5.03 - Bioelectric Measurements: Magnetoencephalography. In *Comprehensive Biomedical Physics*, Anders Brahme (Ed.). Elsevier, Oxford, 47–72. https://doi.org/10.1016/B978-0-444-53632-7.00508-6
[11] Ewan S. Nurse, Philippa J. Karoly, David B. Grayden, and Dean R. Freestone. 2015. A Generalizable Brain-Computer Interface (BCI) Using Machine Learning for Feature Discovery. *PLOS ONE* 10, 6 (June 2015), e0131328. https://doi.org/10.1371/journal.pone.0131328
[12] Dale Purves. 2001. *Neuroscience*. Sinauer Associates, Sunderland, Mass.
[13] Immo Schuetz, T. Scott Murdison, Kevin J. MacKenzie, and Marina Zannoli. 2019. An Explanation of Fitts' Law-like Performance in Gaze-Based Selection Tasks Using a Psychophysics Approach. (May 2019). https://doi.org/10.1145/3290605.3300765
[14] Janaki Sheth, Ariel Tankus, Michelle Tran, Nader Pouratian, Itzhak Fried, and William Speier. 2019. Translating neural signals to text using a Brain-Machine Interface. *arXiv preprint arXiv:1907.04265* (2019).
[15] Eimear Smith and Mark Delargy. 2005. Locked-in syndrome. *BMJ* 330, 7488 (Feb. 2005), 406–409. https://doi.org/10.1136/bmj.330.7488.406
[16] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), Vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
[17] David J. Ward and David J. C. MacKay. 2002. Fast hands-free writing by gaze direction. *Nature* 418, 6900 (Aug. 2002), 838–838. https://doi.org/10.1038/418838a
[18] Francis R. Willett, Donald T. Avansino, Leigh R. Hochberg, Jaimie M. Henderson, and Krishna V. Shenoy. 2021. High-performance brain-to-text communication via handwriting. *Nature* 593, 7858 (May 2021), 249–254. https://doi.org/10.1038/s41586-021-03506-2
[19] Shumin Zhai. 2003. What's in the Eyes for Attentive Input. *Commun. ACM* 46, 3 (mar 2003), 34–39. https://doi.org/10.1145/636772.636795

Requirement ID	Verification Method	Success Criteria
Operations Process:		
A	Inspection	All components connected to electric power source
B	Demonstration	On/off switch functions as expected
C	Demonstration	No output is generated when in Standby Mode
D	Demonstration	Outputs are always generated when in Active mode
E	Test, Analysis	Activating the device and then unplugging the EEG sensors causes a 'Cannot interpret input' error message to be produced on the screen
Operations Performance:		
A	Inspection	Sensors are not invasive
B	Inspection, Test	When subject to reasonable force (as expected in normal use), the sensors do not move more than ±5mm within the containing material
C	Analysis	When timescales for the time to train are recorded on a set of users, less than 5% of users take longer than 30 hours
D	Inspection	The sensors will not be visible from outside the cap
E	Inspection	The device will weigh <3kg
F	Inspection	The cap will weigh <0.5kg
G	Test	The temperature of the device, when measured over a prolonged period of use, does not go outside -40 and 85 degrees Celsius
H	Inspection	The processing unit will be less than 5x4x10 cm
I	Analysis	Over a set of users, an accurate recording of average CER is <10%
J	Analysis	Over a set of users, an accurate recording of average CPM is >80
K	Inspection	The graphic display will be within the areas specified
L	Test	Measured volume during speech generation is between 10 and 85 dB at minimum and maximum volume
M	Analysis	Over a set of >10 users, there is <5% statistical significance that the battery longevity is >4 hours
Operations Safety:		
A	Inspection	The device will have passed MHRA regulation checks
B	Inspection	The device will have passed Regulation 2017/745 regulation checks
C	Inspection	The device will have passed CDRH regulation checks
D	Inspection	The device will have passed EN 60601 regulation checks
E	Inspection	Directives in place to regulate the access and distribution of user data
F	Inspection	Storage and use of user data complies with the Data Protection Act
Operations Documentation:		
A	Inspection	A manual exists containing operating information
Maintenance Process:		
A	Analysis	<5% statistical significance found for a longer timescale of model degradation compared to the baseline of normal neural plasticity timescales on >10 users
Maintenance Performance:		
A	Inspection	A warranty exists that has a period of 10 years, with well-defined warranty conditions
B	Inspection	Directives in place to offer next day support, with well-defined meaning of the phrase 'where reasonable' (e.g. if the user is within 50 miles of a technician)
C	Analysis	Upon significance testing of the days before retraining on a user group of >10 users, the timescale is >30 days with 5% significance
Manufacturing Process:		
A	Inspection	Every product has been pass/fail tested before mass-production phase
B	Inspection	Every 1 in 10 products have been pass/fail tested in mass-production phase
Manufacturing Cost:		
A	Inspection	Products cost less than £5,000 per unit in materials and manufacture before mass production
B	Inspection	Products cost less than £2,000 per unit in materials and manufacture once in mass production
C	Inspection	Device assembly/manufacture costs are less than £10 per unit in mass production
D	Inspection	A 10% held-back store of spare parts is kept, with dynamic reassessment as product distribution rates vary
Installation Process:		
A	Demonstration	Functionality is checked by a trained individual at every device installation
Installation Performance:		
A	Demonstration	Demonstrated functionality of the fasteners between the processing units and output devices to a user's wheelchair
B	Inspection	Directives in place to ensure that a trained individual assists the user in calibrating the device at first installation

Table 2: VCRM for how to verify each of the requirements of Section 2.2.